Finding Hyperplane in Support Vector Machine Problems

by Suneesh Jacob A.

In Support Vector Machine problems, we need to find a boundary hyperplane such that it separates the points into two classes with maximum margin. Let us assume hard margin here, as the points from Figure 1 can be observed to be easily separable. Since there is a free parameter, we also do the scaling. The constrained optimisation problem is

Maximise the margin

$$r$$

subject to

$$y^{(k)}\left(\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{k})})+b\right)\geq r$$

$$\|\boldsymbol{w}\|=1$$

This is equivalent to scaling the data, such that the margin is unity. The equivalent optimisation problem would be

Minimise

$$\frac{1}{2}\|\boldsymbol{w}\|^2=\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}$$

subject to

$$-y^{(k)}\left(\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{k})})+b\right)+1\leq 0 \quad \text{for all } k$$

It can be noted that that $\boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w}=\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x})$ because the quantity is a scalar.

The support vectors are the active constraints, and the remaining constraints will be inactive constraints. Active constraints are treated as equality constraints and inactive constraints are omitted. The Lagrange multipliers corresponding to active constraints should be greater than or equal to zero.

$$L(\boldsymbol{w},b,\boldsymbol{\alpha})=\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}+\sum_j\alpha_j\left(-y^{(j)}\left(\boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w}+b\right)+1\right)$$

$$\Rightarrow L(\boldsymbol{w},b,\boldsymbol{\alpha})=\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}-\sum_j\alpha_j y^{(j)}\left(\boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w}+b\right)+\sum_j\alpha_j$$

$$\Rightarrow L(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} - \sum_j \alpha_j y^{(j)} \boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w} - b\sum_j \alpha_j y^{(j)} + \sum_j \alpha_j$$

Here, $\boldsymbol{\alpha} = \begin{Bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_k\end{Bmatrix}^T$ is the vector of the Lagrange multipliers. If we set the gradient with respect to $\boldsymbol{w}$ and b equal to zero, we get

$$\begin{aligned}
\frac{\partial L}{\partial \boldsymbol{w}} = 0 \Rightarrow \boldsymbol{w}^T - \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}^T(\boldsymbol{x}^{(\boldsymbol{k})}) = 0 \\
\Rightarrow \boldsymbol{w}^T = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}^T(\boldsymbol{x}^{(\boldsymbol{k})}) \\
\Rightarrow \boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{k})})
\end{aligned} \tag{1}$$

and

$$\begin{aligned}
\frac{\partial L}{\partial b} = 0 \Rightarrow -\sum_k \alpha_k y^{(k)} = 0 \\
\Rightarrow \sum_k \alpha_k y^{(k)} = 0
\end{aligned} \tag{2}$$

By substituting these in the Lagrangian, we can derive the dual function as shown below.

$$L(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} - \sum_j \alpha_j y^{(j)} \boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w} - b\sum_j \alpha_j y^{(j)} + \sum_j \alpha_j$$

From Equation (2) we know that $\sum_j \alpha_j y^{(j)} = 0$.

$$\Rightarrow L(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} - \sum_j \alpha_j y^{(j)} \boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w} - b\,(0) + \sum_j \alpha_j$$

$$\Rightarrow L(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} - \sum_j \alpha_j y^{(j)} \boldsymbol{\phi}(\boldsymbol{x})^T\boldsymbol{w} + \sum_j \alpha_j$$

If we substitute $\boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{k})})$, we get

$$\Rightarrow D(\boldsymbol{\alpha}) = L\left(\boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x^{(k)}}), b, \boldsymbol{\alpha}\right) = \frac{1}{2}\left(\sum_i \alpha_i y^{(i)} \boldsymbol{\phi}^T(\boldsymbol{x^{(i)}})\right)\left(\sum_j \alpha_j y^{(j)} \boldsymbol{\phi}(\boldsymbol{x^{(j)}})\right) - \sum_j \alpha_j y^{(j)} \boldsymbol{\phi}(\boldsymbol{x^{(j)}})^T \left(\sum_i \alpha_i y^{(i)} \boldsymbol{\phi}(\boldsymbol{x^{(i)}})\right) + \sum_j \alpha_j$$

$$\Rightarrow D(\boldsymbol{\alpha}) = \frac{1}{2}\left(\sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(i)}}) \boldsymbol{\phi}(\boldsymbol{x^{(j)}})\right) - \left(\sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(i)}}) \boldsymbol{\phi}(\boldsymbol{x^{(j)}})\right) + \sum_j \alpha_j$$

$$\Rightarrow D(\boldsymbol{\alpha}) = -\frac{1}{2}\left(\sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(i)}}) \boldsymbol{\phi}(\boldsymbol{x^{(j)}})\right) + \sum_j \alpha_j$$

Even though this appears to be the dual function, the Lagrange multipliers, i.e., α_i values, are not all independent, because Equation (2) needs to be satisfied - from which one of the Lagrange multipliers can be eliminated. But we will not prefer to do so, because we do not want to lose the symmetry. Hence, we retain the equation as it is. Further more, since the α_i values are Lagrange multipliers corresponding to inequality constraints, they should be greater than or equal to zero. Thus, the dual optimisation problem would become

Maximise

$$D(\boldsymbol{\alpha}) = -\frac{1}{2}\left(\sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(i)}}) \boldsymbol{\phi}(\boldsymbol{x^{(j)}})\right) + \sum_j \alpha_j$$

subject to

$$\sum_k \alpha_k y^{(k)} = 0$$

$$\alpha_k \geq 0$$

Let us keep the $\alpha_k \geq 0$ constraint aside, as we will get back to it later and focus on solving the problem now. The optimisation problem can thus be rewritten as

Minimise

$$D(\boldsymbol{\alpha}) = \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(i)}}) \boldsymbol{\phi}(\boldsymbol{x^{(j)}}) - \sum_j \alpha_j$$

subject to

$$\sum_k \alpha_k y^{(k)} = 0$$

Without breaking the symmetry, let us solve this optimisation problem using Lagrange approach. We would then have the new Lagrangian function as shown below, where μ is the Lagrange parameter for the constraint.

$$\mathcal{L}(\boldsymbol{\alpha}, \mu) = \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) - \sum_j \alpha_j + \mu \sum_j \alpha_j y^{(j)}$$

The solution should satisfy the first order condition that the gradient of the Lagrangian should be zero. This implies $\frac{\partial \mathcal{L}}{\partial \alpha_k} = 0$ for all k and $\frac{\partial \mathcal{L}}{\partial \mu} = 0$.

$$\frac{\partial \mathcal{L}}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k} \left(\frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) - \sum_j \alpha_j + \mu \sum_j \alpha_j y^{(j)} \right)$$

$$\Rightarrow \frac{\partial \mathcal{L}}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k} \left(\frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) \right) - \frac{\partial}{\partial \alpha_k} \left(\sum_j \alpha_j \right) + \frac{\partial}{\partial \alpha_k} \left(\mu \sum_j \alpha_j y^{(j)} \right)$$

$$\Rightarrow \frac{\partial \mathcal{L}}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k} \left(\frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) \right) - \frac{\partial}{\partial \alpha_k} \left(\sum_j \alpha_j \right) + \frac{\partial}{\partial \alpha_k} \left(\mu \sum_j \alpha_j y^{(j)} \right)$$

$$\Rightarrow \frac{\partial \mathcal{L}}{\partial \alpha_i} = \left(\sum_j \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) \right) - 1 + \mu \left(y^{(i)} \right)$$

Hence, $\frac{\partial \mathcal{L}}{\partial \alpha_i} = 0$ gives

$$\left(\sum_j \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) \right) - 1 + \mu \left(y^{(i)} \right) = 0$$

$$\Rightarrow \left(\sum_j \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi}^{\boldsymbol{T}}(\boldsymbol{x}^{(\boldsymbol{i})}) \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{j})}) \right) + \mu y^{(i)} = 1 \tag{3}$$

Likewise,

$$\frac{\partial \mathcal{L}}{\partial \mu} = \sum_j \alpha_j y^{(j)}$$

Hence, $\frac{\partial \mathcal{L}}{\partial \mu} = 0$ gives

$$\sum_j \alpha_j y^{(j)} = 0 \tag{4}$$

Now, let us bring in the fact that α_i values are the Lagrangian multipliers corresponding to inequality constraints, which implies that they should be greater than or equal to zero.

$$\alpha_i \geq 0 \tag{5}$$

Thus, whatever $\boldsymbol{\alpha}$ vector satisfies Equation (3), Equation (4) and Equation (5), would be the solution of the optimisation problem. The equations (3) and (4) would form a system of linear equations in $\boldsymbol{\alpha}$ and μ. Hence, by solving the system of equations, the values $\boldsymbol{\alpha}$ and μ can be found. Once $\boldsymbol{\alpha}$ is found, one could find the weight vector $\boldsymbol{w}$ can be found from Equation (1) as shown below.

$$\boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x^{(k)}})$$

Now, how to compute the bias value b? The value of μ would be the same as the value of b for these SVM problems. The proof for the same is shown below.

From Equation (3), we can have

$$\left(\sum_j \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(i)}}) \boldsymbol{\phi}(\boldsymbol{x^{(j)}}) \right) + \mu y^{(i)} = 1$$

$$\Rightarrow \left(\sum_j \alpha_j y^{(i)} y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(j)}}) \boldsymbol{\phi}(\boldsymbol{x^{(i)}}) \right) + \mu y^{(i)} = 1$$

$$\Rightarrow y^{(i)} \left(\sum_j \alpha_j y^{(j)} \boldsymbol{\phi^T}(\boldsymbol{x^{(j)}}) \right) \boldsymbol{\phi}(\boldsymbol{x^{(i)}}) + \mu y^{(i)} = 1$$

$$\Rightarrow y^{(i)} \left(\boldsymbol{w}^T \right) \boldsymbol{\phi}(\boldsymbol{x^{(i)}}) + \mu y^{(i)} = 1$$

$$\Rightarrow y^{(i)}\left(\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x^{(i)}})+\mu\right)=1$$

$$\Rightarrow \boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x^{(i)}})+\mu=\frac{1}{y^{(i)}}$$

$$\Rightarrow \boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x^{(i)}})=\frac{1}{y^{(i)}}-\mu \tag{6}$$

Now, we know that for any support vector $\left(\boldsymbol{x^{(i)}},y^{(i)}\right)$, we can have

$$y^{(i)}=\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x^{(i)}})+b$$

$$\Rightarrow b=y^{(i)}-\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x^{(i)}})$$

By using Equation (6), we can have

$$b=y^{(i)}-\left(\frac{1}{y^{(i)}}-\mu\right)$$

$$\Rightarrow b=\mu+\left(y^{(i)}-\frac{1}{y^{(i)}}\right)$$

$$\Rightarrow b=\mu+\left(\frac{\left(y^{(i)}\right)^2-1}{y^{(i)}}\right)$$

We know that $y^{(i)}$ is either +1 or -1. This means $\left(y^{(i)}\right)^2=1$ or $\left(y^{(i)}\right)^2-1=0$
Therefore, we can have

$$b=\mu+\left(\frac{\left(y^{(i)}\right)^2-1}{y^{(i)}}\right)$$

$$\Rightarrow b=\mu+\left(\frac{0}{y^{(i)}}\right)$$

$$\Rightarrow b=\mu$$

.

Thus, the method of finding the boundary hyperplane is summarised below:
If $\boldsymbol{\phi}(\boldsymbol{x^{(1)}})$, $\boldsymbol{\phi}(\boldsymbol{x^{(2)}})$, ..., $\boldsymbol{\phi}(\boldsymbol{x^{(n)}})$ are n support vectors of an SVM problem (with hard margin), then

$$\left(y^{(1)}\right)^2 \boldsymbol{\phi^T}(\boldsymbol{x^{(1)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}})\alpha_1 + y^{(1)}y^{(2)}\boldsymbol{\phi^T}(\boldsymbol{x^{(1)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}})\alpha_2 + y^{(1)}y^{(3)}\boldsymbol{\phi^T}(\boldsymbol{x^{(1)}})\boldsymbol{\phi}(\boldsymbol{x^{(3)}})\alpha_3 + \cdots + y^{(1)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(1)}})\boldsymbol{\phi}(\boldsymbol{x^{(n)}})\alpha_n + y^{(1)}\mu = 1$$

$$y^{(2)}y^{(1)}\boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}})\alpha_1 + \left(y^{(2)}\right)^2 \boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}})\alpha_2 + y^{(2)}y^{(3)}\boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(3)}})\alpha_3 + \cdots + y^{(2)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(n)}})\alpha_n + y^{(2)}\mu = 1$$

$$\vdots$$

$$y^{(n)}y^{(1)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}})\alpha_1 + + y^{(n)}y^{(2)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}})\alpha_2 + + y^{(n)}y^{(3)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(3)}})\alpha_2 + \cdots + \left(y^{(n)}\right)^2 \boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(n)}})\alpha_n + y^{(n)}\mu = 1$$

$$\alpha_1 y^{(1)} + \alpha_2 y^{(2)} + \alpha_3 y^{(3)} + \cdots + \alpha_n y^{(n)} = 0$$

with $\alpha_i \geq 0$.
In concise form, it can be written as

$$\begin{bmatrix} [\boldsymbol{A}] & \{\boldsymbol{y}\} \\ \{\boldsymbol{y}\}^T & 0 \end{bmatrix} \begin{bmatrix} \{\boldsymbol{\alpha}\} \\ \mu \end{bmatrix} = \begin{Bmatrix} \{\boldsymbol{1}\} \\ 0 \end{Bmatrix} = \begin{Bmatrix} 1 \\ \vdots \\ 1 \\ 0 \end{Bmatrix}$$

where

$$[\boldsymbol{A}] = \begin{bmatrix} \left(y^{(1)}\right)^2 \boldsymbol{\phi^T}(\boldsymbol{x^{(1)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}}) & y^{(1)}y^{(2)}\boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}}) & y^{(1)}y^{(3)}\boldsymbol{\phi^T}(\boldsymbol{x^{(3)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}}) & \cdots & y^{(1)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}}) \\ y^{(1)}y^{(2)}\boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}}) & \left(y^{(2)}\right)^2 \boldsymbol{\phi^T}(\boldsymbol{x^{(2)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}}) & y^{(2)}y^{(3)}\boldsymbol{\phi^T}(\boldsymbol{x^{(3)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}}) & \cdots & y^{(2)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y^{(1)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(1)}}) & y^{(2)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(2)}}) & y^{(3)}y^{(n)}\boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(3)}}) & \cdots & \left(y^{(n)}\right)^2 \boldsymbol{\phi^T}(\boldsymbol{x^{(n)}})\boldsymbol{\phi}(\boldsymbol{x^{(n)}}) \end{bmatrix}$$

$$\{\boldsymbol{y}\} = \begin{Bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \\ \vdots \\ y^{(n)} \end{Bmatrix}$$

Provided that all the values α_i are greater than or equal to zero.

Problem

For the given dataset below, construct a linear hyperplane in Support Vector Machine

S.No	x_1	x_2	y
1	4	1	-1
2	2	4	-1
3	2	3	-1
4	3	6	-1
5	4	4	-1
6	9	10	1
7	6	8	1
8	9	5	1
9	8	7	1
10	10	8	1

Solution

From Figure 1, we can find the support vectors as (3,6) and (6,4) with labels as -1 and +1, respectively. Hence, we have

$$\boldsymbol{\phi}\left(\boldsymbol{x^{(1)}}\right) = \begin{Bmatrix} 3 \\ 6 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) = \begin{Bmatrix} 6 \\ 8 \end{Bmatrix}$$

$$y^{(1)} = -1$$

$$y^{(2)} = +1$$

Therefore, we can have

$$[A] = \begin{bmatrix} \left(y^{(1)}\right)^2 \boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(1)}}\right) & y^{(1)}y^{(2)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) \\ y^{(1)}y^{(2)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) & \left(y^{(2)}\right)^2 \boldsymbol{\phi^T}\left(\boldsymbol{x^{(2)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) \end{bmatrix} = \begin{bmatrix} (-1)^2\,(3\times 3+6\times 6) & (-1)(+1)\,(3\times 6+6\times 8) \\ (-1)(+1)\,(3\times 6+6\times 8) & (+1)^2\,(6\times 6+8\times 8) \end{bmatrix} = \begin{bmatrix} 45 & -66 \\ -66 & 100 \end{bmatrix}$$

$$\{\boldsymbol{y}\} = \begin{Bmatrix} y^{(1)} \\ y^{(2)} \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

Hence, we can have

$$\begin{bmatrix} [\boldsymbol{A}] & \{\boldsymbol{y}\} \\ \{\boldsymbol{y}\}^T & 0 \end{bmatrix} \begin{bmatrix} \{\boldsymbol{\alpha}\} \\ \mu \end{bmatrix} = \begin{Bmatrix} \{\boldsymbol{1}\} \\ 0 \end{Bmatrix}$$

$$\begin{bmatrix} \begin{bmatrix} 45 & -66 \\ -66 & 100 \end{bmatrix} & \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} \\ \begin{Bmatrix} -1 & 1 \end{Bmatrix} & 0 \end{bmatrix} \begin{bmatrix} \begin{Bmatrix} \alpha_1 \\ \alpha_2 \end{Bmatrix} \\ \mu \end{bmatrix} = \begin{Bmatrix} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \\ 0 \end{Bmatrix}$$

$$\Rightarrow \begin{bmatrix} 45 & -66 & -1 \\ -66 & 100 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \mu \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \\ 0 \end{Bmatrix}$$

$$\Rightarrow \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \mu \end{Bmatrix} = \begin{bmatrix} 45 & -66 & -1 \\ -66 & 100 & 1 \\ -1 & 1 & 0 \end{bmatrix}^{-1} \begin{Bmatrix} 1 \\ 1 \\ 0 \end{Bmatrix} = \begin{Bmatrix} \frac{2}{13} \\ \frac{2}{13} \\ -\frac{55}{13} \end{Bmatrix}$$

Now, we know that

$$\boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x^{(k)}}) = \alpha_1 y^{(1)} \boldsymbol{\phi}(\boldsymbol{x^{(1)}}) + \alpha_2 y^{(2)} \boldsymbol{\phi}(\boldsymbol{x^{(2)}}) = \frac{2}{13} \times (-1) \times \begin{Bmatrix} 3 \\ 6 \end{Bmatrix} + \frac{2}{13} \times (+1) \times \begin{Bmatrix} 6 \\ 8 \end{Bmatrix} = \begin{Bmatrix} \frac{6}{13} \\ \frac{4}{13} \end{Bmatrix}$$

and

$$b = \mu = -\frac{55}{13}$$

Therefore the boundary line is

$$\boldsymbol{w}^T \boldsymbol{\phi}(\boldsymbol{x}) + b = 0$$

$$\Rightarrow w_1 x_1 + w_2 x_2 + b = 0$$

$$\Rightarrow \frac{6}{13} x_1 + \frac{4}{13} x_2 - \frac{55}{13} = 0$$

$$\Rightarrow 6x_1 + 4x_2 - 55 = 0$$

Problem 2

For the given dataset below, construct a linear hyperplane in Support Vector Machine

S.No	x_1	x_2	y
1	3	1	1
2	3	-1	1
3	6	1	1
4	6	-1	1
5	1	0	-1
6	0	1	-1
7	0	-1	-1
8	-1	0	-1

Solution

From Figure 3, we can find the support vectors as (1,0), (3,1) and (3,-1) with labels as -1, +1 and +1, respectively. Hence, we have

$$\boldsymbol{\phi}\left(\boldsymbol{x^{(1)}}\right) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) = \begin{Bmatrix} 3 \\ 1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x^{(3)}}\right) = \begin{Bmatrix} 3 \\ -1 \end{Bmatrix}$$

$$y^{(1)} = -1$$

$$y^{(2)} = +1$$

$$y^{(3)} = +1$$

Therefore, we can have

$$[A] = \begin{bmatrix} \left(y^{(1)}\right)^2 \boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(1)}}\right) & y^{(1)}y^{(2)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) & y^{(1)}y^{(3)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(3)}}\right) \\ y^{(1)}y^{(2)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) & \left(y^{(2)}\right)^2 \boldsymbol{\phi^T}\left(\boldsymbol{x^{(2)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(2)}}\right) & y^{(2)}y^{(3)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(2)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(3)}}\right) \\ y^{(1)}y^{(3)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(1)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(3)}}\right) & y^{(2)}y^{(3)}\boldsymbol{\phi^T}\left(\boldsymbol{x^{(2)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(3)}}\right) & \left(y^{(3)}\right)^2 \boldsymbol{\phi^T}\left(\boldsymbol{x^{(3)}}\right)\boldsymbol{\phi}\left(\boldsymbol{x^{(3)}}\right) \end{bmatrix}$$

$$= \begin{bmatrix} (-1)^2\,(1\times 1 + 0\times 0) & (-1)(+1)\,(1\times 3 + 0\times 1) & (-1)(+1)\,(1\times 3 + 0\times -1) \\ (-1)(+1)\,(1\times 3 + 0\times 1) & (+1)^2\,(3\times 3 + 1\times 1) & (+1)(+1)\,(3\times 3 + 1\times -1) \\ (-1)(+1)\,(1\times 3 + 0\times -1) & (+1)(+1)\,(3\times 3 + 1\times -1) & (+1)^2\,(3\times 3 + (-1)\times(-1)) \end{bmatrix} = \begin{bmatrix} 1 & -3 & -3 \\ -3 & 10 & 8 \\ -3 & 8 & 10 \end{bmatrix}$$

$$\{\boldsymbol{y}\} = \begin{Bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \\ 1 \end{Bmatrix}$$

Hence, we can have

$$\begin{bmatrix} [\boldsymbol{A}] & \{\boldsymbol{y}\} \\ \{\boldsymbol{y}\}^T & 0 \end{bmatrix} \begin{bmatrix} \{\boldsymbol{\alpha}\} \\ \mu \end{bmatrix} = \begin{Bmatrix} \{\boldsymbol{1}\} \\ 0 \end{Bmatrix}$$

$$\left[\begin{bmatrix} 1 & -3 & -3 \\ -3 & 10 & 8 \\ -3 & 8 & 10 \\ \{-1 & 1 & 1\} \end{bmatrix} \quad \begin{Bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{Bmatrix}\right] \left[\begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{Bmatrix} \atop \mu\right] = \left\{\begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \atop 0\right\}$$

$$\Rightarrow \begin{bmatrix} 1 & -3 & -3 & -1 \\ -3 & 10 & 8 & 1 \\ -3 & 8 & 10 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix} \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \mu \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{Bmatrix}$$

$$\Rightarrow \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \mu \end{Bmatrix} = \begin{bmatrix} 1 & -3 & -3 & -1 \\ -3 & 10 & 8 & 1 \\ -3 & 8 & 10 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix}^{-1} \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{Bmatrix} = \begin{Bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \\ -2 \end{Bmatrix}$$

Now, we know that

$$\boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{k})}) = \alpha_1 y^{(1)} \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{1})}) + \alpha_2 y^{(2)} \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{2})}) + \alpha_3 y^{(3)} \boldsymbol{\phi}(\boldsymbol{x}^{(\boldsymbol{3})}) = \frac{1}{2} \times (-1) \times \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} + \frac{1}{4} \times (+1) \times \begin{Bmatrix} 3 \\ 1 \end{Bmatrix} + \frac{1}{4} \times (+1) \times \begin{Bmatrix} 3 \\ -1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$$

and

$$b = \mu = -2$$

Therefore the boundary line is

$$\boldsymbol{w}^T \boldsymbol{\phi}\left(\boldsymbol{x}\right) + b = 0$$

$$\Rightarrow w_1 x_1 + w_2 x_2 + b = 0$$

$$\Rightarrow (1)x_1 + (0)x_2 - 2 = 0$$

$$\Rightarrow x_1 - 2 = 0$$

The boundary

Problem 3

For the given dataset below, construct a linear hyperplane in Support Vector Machine, by choosing the basis functions as $\boldsymbol{\phi}(\boldsymbol{x}) = \begin{Bmatrix} x_1 \\ x_2 \\ \frac{(x_1^2+x_2^2)-5}{3} \end{Bmatrix}$.

S.No	x_1	x_2	y
1	2	2	1
2	2	-2	1
3	-2	-2	1
4	-2	2	1
5	1	1	-1
6	1	-1	-1
7	-1	-1	-1
8	-1	1	-1

Solution

If we project the points to the basis space $\boldsymbol{\phi}(\boldsymbol{x}) = \begin{Bmatrix} x_1 \\ x_2 \\ \frac{(x_1^2+x_2^2)-5}{3} \end{Bmatrix}$, we get the below points:

S.No	$\phi_1(\boldsymbol{x}) = x_1$	$\phi_2(\boldsymbol{x}) = x_2$	$\phi_3(\boldsymbol{x}) = \frac{(x_1^2+x_2^2)-5}{3}$	y
1	2	2	1	1
2	2	-2	1	1
3	-2	-2	1	1
4	-2	2	1	1
5	1	1	-1	-1
6	1	-1	-1	-1
7	-1	-1	-1	-1
8	-1	1	-1	-1

The points are shown in Figure 5.
From Figure 6, we can find that all the points are support vectors. Hence, we have

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(1)}\right) = \begin{Bmatrix} 2 \\ 2 \\ 1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(2)}\right) = \begin{Bmatrix} 2 \\ -2 \\ 1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(3)}\right) = \begin{Bmatrix} -2 \\ -2 \\ 1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(4)}\right) = \begin{Bmatrix} -2 \\ 2 \\ 1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(5)}\right) = \begin{Bmatrix} 1 \\ 1 \\ -1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(6)}\right) = \begin{Bmatrix} 1 \\ -1 \\ -1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(7)}\right) = \begin{Bmatrix} -1 \\ -1 \\ -1 \end{Bmatrix}$$

$$\boldsymbol{\phi}\left(\boldsymbol{x}^{(8)}\right) = \begin{Bmatrix} -1 \\ 1 \\ -1 \end{Bmatrix}$$

$$y^{(1)} = +1$$

$$y^{(2)} = +1$$

$$y^{(3)} = +1$$

$$y^{(4)} = +1$$

$$y^{(5)} = -1$$

$$y^{(6)} = -1$$

$$y^{(7)} = -1$$

$$y^{(8)} = -1$$

Therefore, we can have

$$[A] = \begin{bmatrix} \left(y^{(1)}\right)^2 \boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{1})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{1})}\right) & y^{(1)}y^{(2)}\boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{1})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{2})}\right) & \cdots & y^{(1)}y^{(8)}\boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{1})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{8})}\right) \\ y^{(1)}y^{(2)}\boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{1})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{2})}\right) & \left(y^{(2)}\right)^2 \boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{2})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{2})}\right) & \cdots & y^{(2)}y^{(8)}\boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{2})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{8})}\right) \\ \vdots & \vdots & \ddots & \vdots \\ y^{(1)}y^{(8)}\boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{1})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{8})}\right) & y^{(2)}y^{(8)}\boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{2})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{8})}\right) & \cdots & \left(y^{(8)}\right)^2 \boldsymbol{\phi}^{\boldsymbol{T}}\left(\boldsymbol{x}^{(\mathbf{8})}\right)\boldsymbol{\phi}\left(\boldsymbol{x}^{(\mathbf{8})}\right) \end{bmatrix}$$

$$= \begin{bmatrix} 9 & 1 & -7 & 1 & -3 & 1 & 5 & 1 \\ 1 & 9 & 1 & -7 & 1 & -3 & 1 & 5 \\ -7 & 1 & 9 & 1 & 5 & 1 & -3 & 1 \\ 1 & -7 & 1 & 9 & 1 & 5 & 1 & -3 \\ -3 & 1 & 5 & 1 & 3 & 1 & -1 & 1 \\ 1 & -3 & 1 & 5 & 1 & 3 & 1 & -1 \\ 5 & 1 & -3 & 1 & -1 & 1 & 3 & 1 \\ 1 & 5 & 1 & -3 & 1 & -1 & 1 & 3 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

$$
\{\boldsymbol{y}\} = \begin{Bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \\ y^{(4)} \\ y^{(5)} \\ y^{(6)} \\ y^{(7)} \\ y^{(8)} \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \end{Bmatrix}
$$

Hence, we can have

$$
\begin{bmatrix} [\boldsymbol{A}] & \{\boldsymbol{y}\} \\ \{\boldsymbol{y}\}^T & 0 \end{bmatrix} \begin{bmatrix} \{\boldsymbol{\alpha}\} \\ \mu \end{bmatrix} = \begin{Bmatrix} \{\mathbf{1}\} \\ 0 \end{Bmatrix}
$$

$$
\begin{bmatrix} \begin{bmatrix} 9 & 1 & -7 & 1 & -3 & 1 & 5 & 1 \\ 1 & 9 & 1 & -7 & 1 & -3 & 1 & 5 \\ -7 & 1 & 9 & 1 & 5 & 1 & -3 & 1 \\ 1 & -7 & 1 & 9 & 1 & 5 & 1 & -3 \\ -3 & 1 & 5 & 1 & 3 & 1 & -1 & 1 \\ 1 & -3 & 1 & 5 & 1 & 3 & 1 & -1 \\ 5 & 1 & -3 & 1 & -1 & 1 & 3 & 1 \\ 1 & 5 & 1 & -3 & 1 & -1 & 1 & 3 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix} & \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \end{Bmatrix} \\ \{1 \quad 1 \quad 1 \quad 1 \quad -1 \quad -1 \quad -1 \quad -1\} & 0 \end{bmatrix} \begin{bmatrix} \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{Bmatrix} \\ \mu \end{bmatrix} = \begin{Bmatrix} \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \\ 0 \end{Bmatrix}
$$

$$
\Rightarrow \begin{bmatrix} 9 & 1 & -7 & 1 & -3 & 1 & 5 & 1 & 1 \\ 1 & 9 & 1 & -7 & 1 & -3 & 1 & 5 & 1 \\ -7 & 1 & 9 & 1 & 5 & 1 & -3 & 1 & 1 \\ 1 & -7 & 1 & 9 & 1 & 5 & 1 & -3 & 1 \\ -3 & 1 & 5 & 1 & 3 & 1 & -1 & 1 & -1 \\ 1 & -3 & 1 & 5 & 1 & 3 & 1 & -1 & -1 \\ 5 & 1 & -3 & 1 & -1 & 1 & 3 & 1 & -1 \\ 1 & 5 & 1 & -3 & 1 & -1 & 1 & 3 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 0 \end{bmatrix} \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \mu \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{Bmatrix}
$$

$$\Rightarrow \left\{\begin{matrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \mu\end{matrix}\right\} = \begin{bmatrix} 9 & 1 & -7 & 1 & -3 & 1 & 5 & 1 & 1\\ 1 & 9 & 1 & -7 & 1 & -3 & 1 & 5 & 1\\ -7 & 1 & 9 & 1 & 5 & 1 & -3 & 1 & 1\\ 1 & -7 & 1 & 9 & 1 & 5 & 1 & -3 & 1\\ -3 & 1 & 5 & 1 & 3 & 1 & -1 & 1 & -1\\ 1 & -3 & 1 & 5 & 1 & 3 & 1 & -1 & -1\\ 5 & 1 & -3 & 1 & -1 & 1 & 3 & 1 & -1\\ 1 & 5 & 1 & -3 & 1 & -1 & 1 & 3 & -1\\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 0 \end{bmatrix}^{-1} \left\{\begin{matrix}1\\1\\1\\1\\1\\1\\1\\1\\0\end{matrix}\right\}$$

But, this is a singular matrix and hence cannot be inverted. In fact, the rank of this 9×9 matrix turns out to be just 5. This means, only five independent equations exist. The rest of the equations have to be linearly dependent on the five independent equations. One could reduce it (detailed steps are not being shown here) to the set of equations as shown below.

$$\begin{aligned} \alpha_1 &= -\alpha_4 - \tfrac{1}{2}\alpha_6 - \tfrac{1}{2}\alpha_7 + \tfrac{3}{8}\\ \alpha_2 &= \alpha_4 + \tfrac{1}{2}\alpha_6 - \tfrac{1}{2}\alpha_8 + \tfrac{3}{8}\\ \alpha_3 &= -\alpha_4 + \tfrac{1}{2}\alpha_7 + \tfrac{1}{2}\alpha_8 + \tfrac{1}{8}\\ \alpha_5 &= -\alpha_6 - \alpha_7 - \alpha_8 + \tfrac{1}{2}\\ \mu &= 0 \end{aligned}$$

Here we have five equations for 9 unknowns, because the rank of the matrix is 5. Any solution satisfying these equations should work. For simplicity, let's assume $\alpha_4 = \alpha_6 = \alpha_7 = \alpha_8 = 0$. This would yield $\alpha_1 = \frac{3}{8}$, $\alpha_2 = 0$, $\alpha_3 = \frac{1}{8}$ and $\alpha_5 = \frac{1}{2}$.

Now, we know that

$$\boldsymbol{w} = \sum_k \alpha_k y^{(k)} \boldsymbol{\phi}(\boldsymbol{x^{(k)}}) = \alpha_1 y^{(1)} \boldsymbol{\phi}(\boldsymbol{x^{(1)}}) + \alpha_2 y^{(2)} \boldsymbol{\phi}(\boldsymbol{x^{(2)}}) + \alpha_3 y^{(3)} \boldsymbol{\phi}(\boldsymbol{x^{(3)}}) + \alpha_4 y^{(4)} \boldsymbol{\phi}(\boldsymbol{x^{(4)}}) + \alpha_5 y^{(5)} \boldsymbol{\phi}(\boldsymbol{x^{(5)}}) + \alpha_6 y^{(6)} \boldsymbol{\phi}(\boldsymbol{x^{(6)}}) + \alpha_7 y^{(7)} \boldsymbol{\phi}(\boldsymbol{x^{(7)}}) + \alpha_8 y^{(8)} \boldsymbol{\phi}(\boldsymbol{x^{(8)}})$$

$$\Rightarrow \boldsymbol{w} = \alpha_1 y^{(1)} \boldsymbol{\phi}(\boldsymbol{x^{(1)}}) + \alpha_3 y^{(3)} \boldsymbol{\phi}(\boldsymbol{x^{(3)}}) + \alpha_5 y^{(5)} \boldsymbol{\phi}(\boldsymbol{x^{(5)}})$$

$$\Rightarrow \boldsymbol{w} = \frac{3}{8} \times (+1) \times \left\{\begin{matrix}2\\2\\1\end{matrix}\right\} + \frac{1}{8} \times (+1) \times \left\{\begin{matrix}-2\\-2\\1\end{matrix}\right\} + \frac{1}{2} \times (-1) \times \left\{\begin{matrix}1\\1\\-1\end{matrix}\right\} = \left\{\begin{matrix}0\\0\\1\end{matrix}\right\}$$

and

$$b = \mu = 0$$

Therefore the boundary plane is

$$\boldsymbol{w}^T \boldsymbol{\phi}(\boldsymbol{x}) + b = 0$$

$$\Rightarrow w_1 \phi_1(\boldsymbol{x}) + w_2 \phi_2(\boldsymbol{x}) + w_3 \phi_3(\boldsymbol{x}) + b = 0$$

$$\Rightarrow w_1 x_1 + w_2 x_2 + w_3 \left(\frac{x_1^2 + x_2^2 - 5}{3} \right) + b = 0$$

$$\Rightarrow 0 \times x_1 + \times x_2 + 1 \times \left(\frac{x_1^2 + x_2^2 - 5}{3} \right) + 0 = 0$$

$$\Rightarrow \frac{x_1^2 + x_2^2 - 5}{3} = 0$$

$$\Rightarrow x_1^2 + x_2^2 - 5 = 0$$

The boundary of the hyperplane is shown in Figure 7 and the boundary curve is shown in Figure 8.

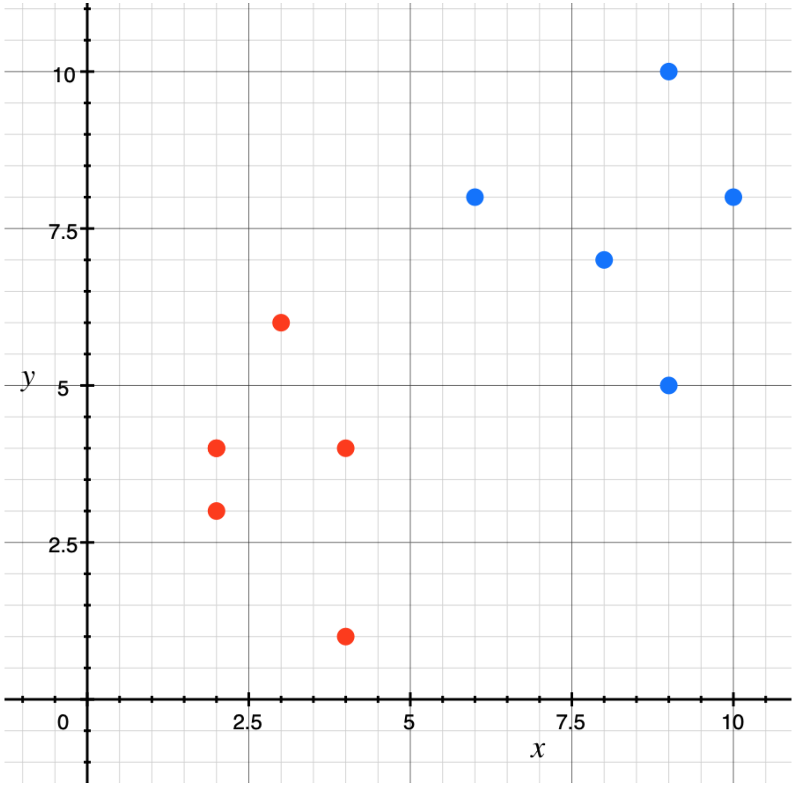

Figure 1: Points

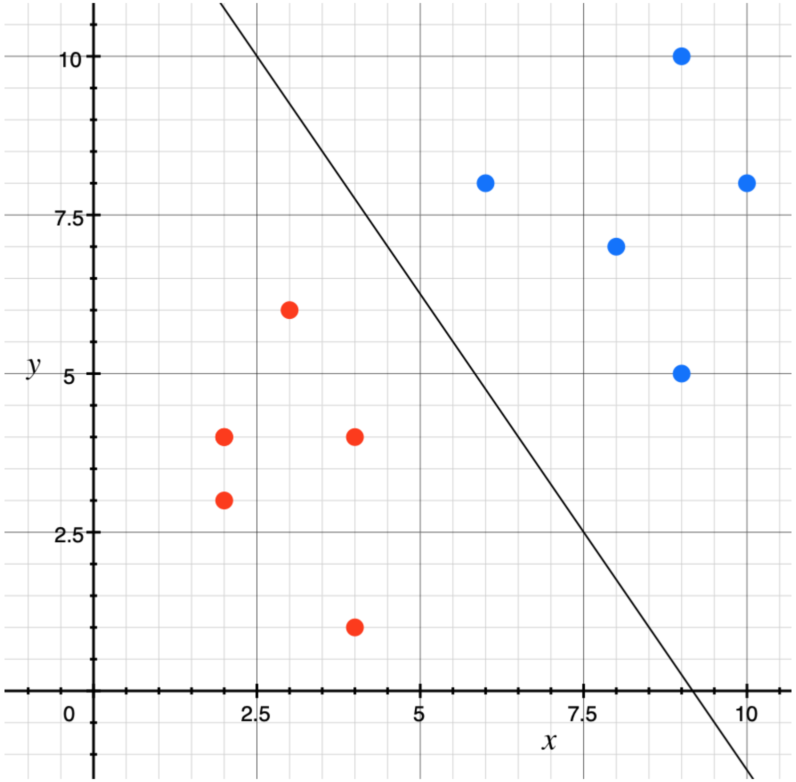

Figure 2: Points with boundary line

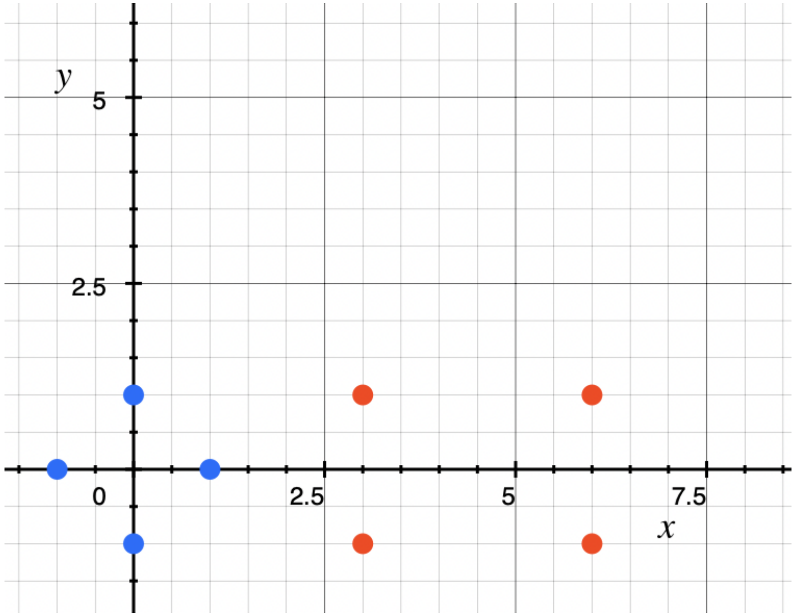

Figure 3: Points

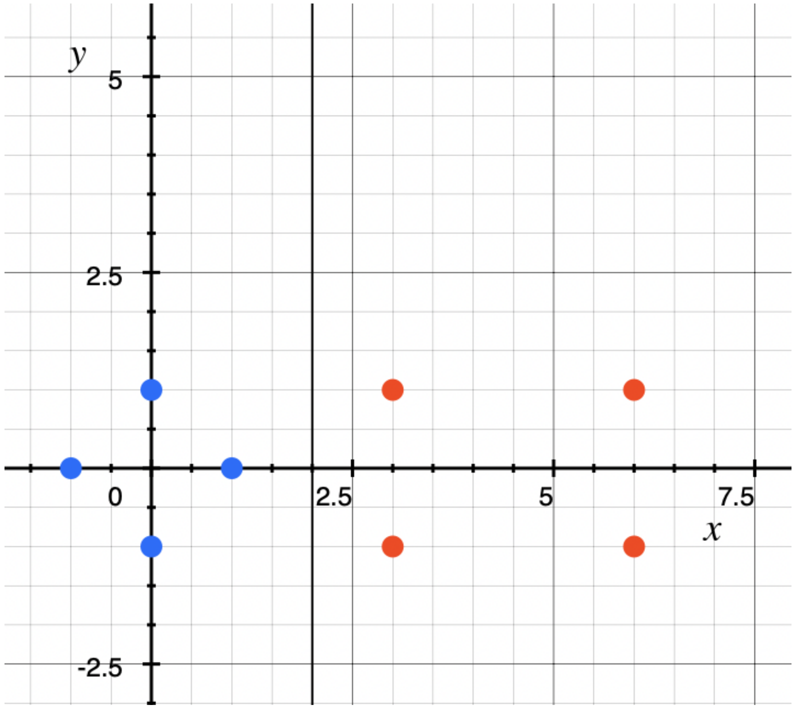

Figure 4: Points with boundary line

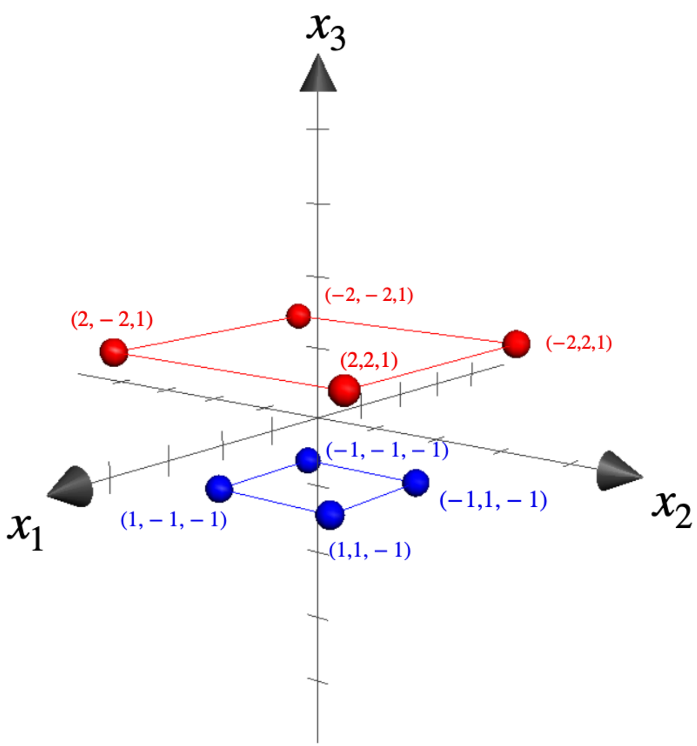

Figure 5: Points in the space of basis functions

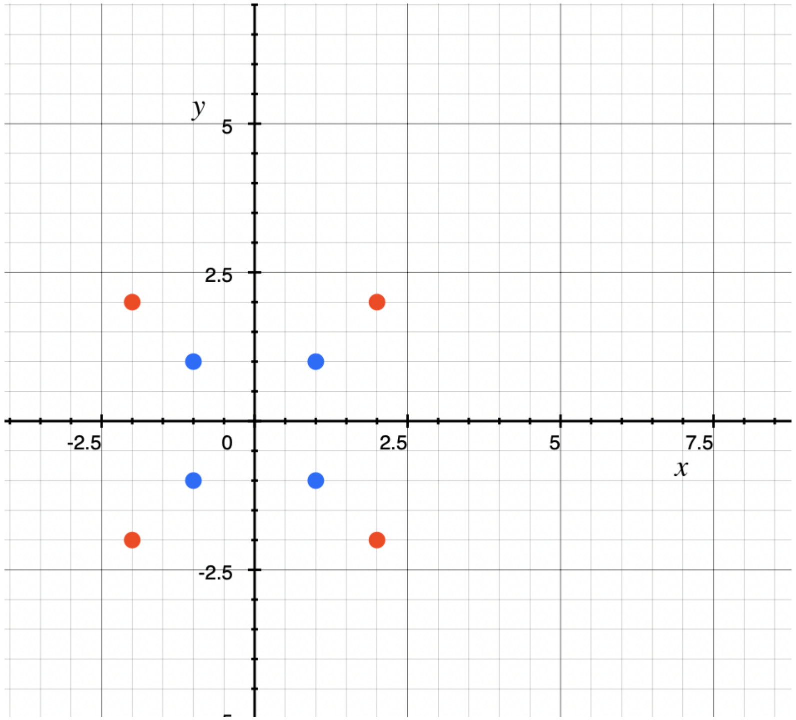

Figure 6: Points

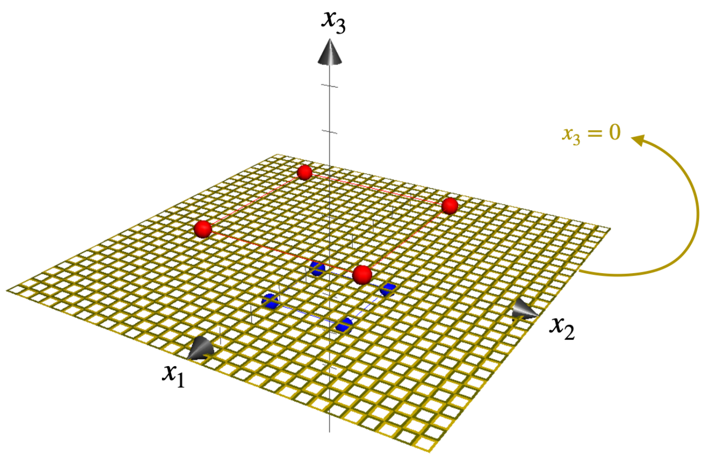

Figure 7: Points with boundary plane

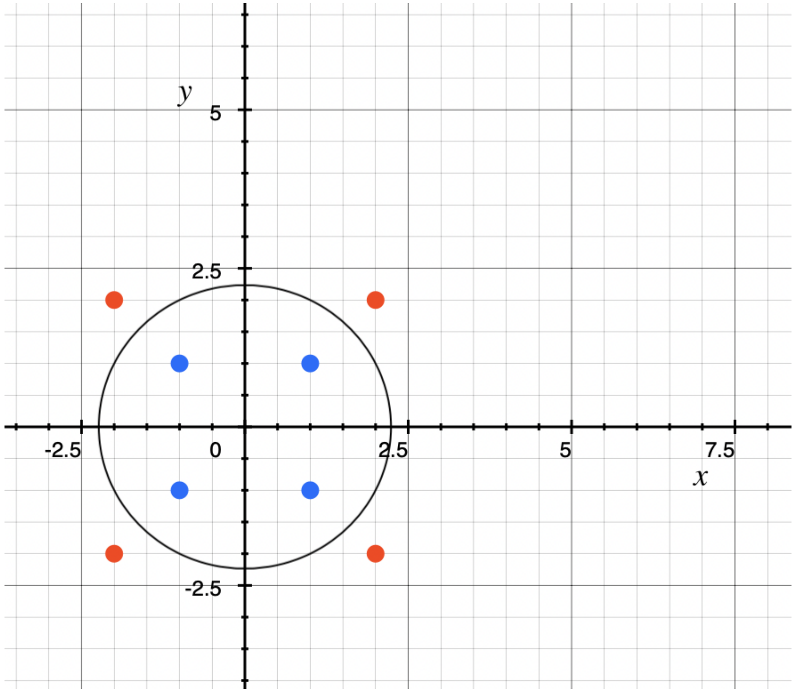

Figure 8: Points with boundary curve